CSP simplification of chemical kinetic systems under uncertainty

Thomas M. K. Coles*, Habib N. Najm†, and Youssef M. Marzouk*
*Massachusetts Institute of Technology; Cambridge, MA USA
†Sandia National Laboratories; Livermore, CA USA

***Abstract*— Chemical kinetic systems contain both considerable uncertainty in their rate parameters and dynamics at multiple time scales. The latter feature aids model reduction in the deterministic case, but model reduction under uncertainty raises new challenges. We use computational singular perturbation (CSP) to calculate probabilistic "importance indices" for species-reaction pairs. Distributions of these indices are used to form reduced models that yield predictions *within* probabilistic bounds determined by the full model, or, alternatively, preserve entire output distributions of the full model.**

I. Overview

Simulation of chemical systems with detailed kinetics can be computationally intensive. While homogeneous systems containing hundreds of species and thousands of elementary reactions can be integrated in reasonable times, a direct approach is not practical when numerous calculations are required, e.g., in the numerical simulation of reacting flow or in problems of optimization or sampling. Here, model reduction techniques are required for computational tractability. Simplified or reduced kinetic models can also provide insight into chemical systems by revealing key pathways and interactions. Many existing methods for reduction in a deterministic setting take advantage of timescale separation in chemical kinetic systems, where different timescales result from a range of slow and fast reactions [1], [2], [3]. However, the relevant reaction rate parameters are typically uncertain and the impact of this uncertainty on model reduction has not previously been investigated.

We use the method of computational singular perturbation (CSP) [4] to calculate probabilistic 'importance indices' for species-reaction pairs on both fast and slow timescales. By modifying an existing deterministic algorithm [5], distributions of these indices are used to generate new reduced models that take account of rate parameter uncertainty. We replace the deterministic threshold used in the original algorithm with a new threshold based on the CVaR (Conditional Value at Risk) [6]. First, a 95% confidence level for each importance index is found; the CVaR is then the expected value of an importance index given that it has exceeded the confidence level. This quantity provides a measure of risk and reduces the impact of both sampling errors and the subjective choice of threshold values.

Different error criteria are used to examine these new reduced models for their ability to either (1) yield predictions within probabilistic bounds determined by the full model, or (2) preserve entire output probability distributions of the full model. Objective (1) allows for greater model reduction by taking advantage of uncertainty to allow larger errors compared to the deterministic case. Objective (2) gives greater confidence in the quality of the reduced model over a wider range of parameters. These objectives do not always conflict; sometimes both objectives can be partially satisfied.

We compare our results with both existing deterministic algorithms and with an exhaustive search for the best reduced model of given size in a methane-air reaction mechanism at nominal rate parameters. Although the deterministic simplification algorithm performs well, it has not previously been compared to optimal results from an exhaustive search. We find that there is an opportunity for considerable improvement to the algorithm in terms of both error in ignition delay at nominal rate parameters and smallest achievable mechanism size. The modified algorithm succeeds in allowing smaller mechanism sizes, but does not give lower error with nominal rate parameters at sizes that can already be achieved. However, it significantly improves reproduction of the output distribution when sampling from a distribution of rate parameters.

II. Numerical examples

A. CSP versus optimal simplified models

We first present results obtained in an entirely deterministic context. Skeletal mechanisms generated using the CSP simplification algorithm of Valorani et al. [5] are compared to "optimal" simplified mechanisms of a given size, found through an exhaustive search of all possible simplified mechanisms. Exhaustive search has combinatorial complexity and is obviously not practical as a simplification algorithm, but the goal of this comparison is to assess how far the results of a fast *heuristic* simplification algorithm lie from the best possible skeletal model. This assessment has additional relevance below, where we consider two new error criteria that incorporate uncertainty, then introduce a *new* heuristic simplification algorithm designed to do a "good" job minimizing these errors.

Figure 1 shows relative error in ignition delay time as predicted by simplified mechanisms of differing sizes (where size corresponds to the number of retained species). The simplified mechanisms are found through CSP simplification or by exhaustive search, as described above. The "full" mechanism in this example is GRIMech

3.0 [7] with C_3 and NO_x chemistry removed. (Removal of these species/reactions is simply a practical choice, intended to make exhaustive search more computationally feasible.) The ignition delay of a hydrogen-enriched methane-air mixture under constant-pressure and adiabatic conditions is calculated for three different initial temperatures and mixture compositions; the figure shows the *relative* error in this ignition delay, averaged over all three cases.

Both the optimal and CSP-produced simplified mechanisms show a trend of smaller errors with larger mechanism sizes. Error of the optimal mechanisms is strictly non-increasing with mechanism size, as expected. For intermediate sizes, the CSP-produced mechanisms yield errors that are within one order of magnitude of the optimal (minimal) error; this performance may be considered quite reasonable, given the vastly different computational costs of the two simplification schemes.

Figure 2 shows the species actually removed in the CSP-simplified model and in the optimal simplified model of any size. Species removed by CSP simplification form nested sets—i.e., if a species is removed in mechanism of size n, it necessarily will be removed in mechanisms of size $n-1$ or smaller, as the latter result from larger values of the importance index threshold parameter. Exhaustive search, on the other hand, has no such constraints; thus, the removed species in an optimal mechanism of a given size do not need to be nested.

It is also worth noting that these results do not fully evaluate the generality of the simplified mechanisms. While the exhaustive-search mechanisms are *optimal for the prescribed error criterion*, it is entirely possible that the CSP-simplified mechanisms might better predict ignition delays at alternate initial conditions, or better reproduce different functionals of the species trajectories. While the heuristic CSP simplification algorithm takes into account the dynamics and structure of the chemical mechanism, it does not optimize for any particular output error criterion.

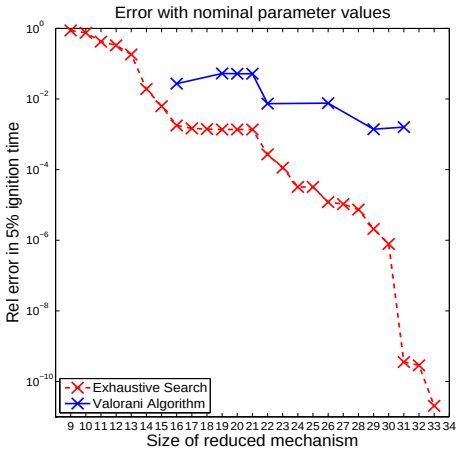

Fig. 1. Relative error in ignition delay predictions, for mechanisms obtained via CSP simplification (blue) and optimal mechanisms obtained via exhaustive search (red).

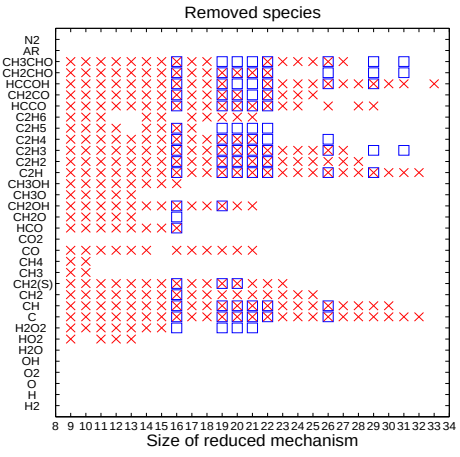

Fig. 2. Species removed by CSP simplification (blue squares) and by exhaustive search for an optimal mechanism of given size (red '×'s).

B. New error criteria for reduction under uncertainty

One of the key notions we wish to explore in the present work is whether uncertainty provides opportunities for model reduction over a deterministic case. In other words, is it possible—and reasonable—to reduce a detailed mechanism that has uncertain parameters *further* than a detailed mechanism with precise/deterministic parameters?

We approach this question by devising output error criteria that will guide model reduction under uncertainty. A first criterion states that the difference between the simplified model prediction y_s and the detailed model prediction y_d of some output quantity of interest should be *normalized* by the standard deviation of the detailed model prediction. The criterion is thus

$$\text{err} = \frac{|y_s - y_d|}{\sigma_d}. \tag{1}$$

A more general criterion comes from the realization that, given a probabilistic description of uncertainty in the rate parameters, both the full-model and simplified-model outputs are random variables. Accordingly, one may choose among many established notions of "distance" between probability measures [8]. Here, we employ the Kullback-Leibler (K-L) divergence. If y_s has probability density p_s and y_d has probability density p_d, the K-L divergence from p_d to p_s is

$$\begin{aligned} D_{KL}\left(p_d \| p_s\right) &= \int p_d(y) \log \frac{p_d(y)}{p_s(y)}\, dy \\ &= \mathbb{E}_{p_d}\left[\log \frac{p_d(y)}{p_s(y)}\right]. \end{aligned} \tag{2}$$

We note that the K-L divergence can reduce to (1) in

the case of Gaussian output distributions with identical variance.

C. CSP simplification with risk-based thresholding

We modify the CSP simplification algorithm to account for parametric uncertainty in the detailed mechanism. CSP-based simplification relies on the evaluation of slow and fast CSP importance indices I^s_{km} and I^r_{km}, which quantify the importance of reaction k to the slow or fast evolution of species m. Uncertainty in rate parameters (e.g., pre-exponential factors, activation energies, temperature exponents, etc) gives rise to uncertainty in these importance indices. Given a probabilistic description of the rate parameter uncertainties, the importance indices are naturally endowed with a joint probability distribution.

The importance indices calculated here are normalized for each species, as suggested in [9]. Additionally, we consider only the time average of each normalized importance index; averages are calculated over the time trajectory of each ignition process. Monte Carlo simulation of homogeneous autoignition, over the space of uncertain rate parameters, is used to construct the probability distribution of the importance indices. (Polynomial chaos approximations [10] to the importance indices are, of course, a viable and potentially far more efficient alternative to Monte Carlo; we shall explore this avenue in future work.) We consider eight uncertain parameters; all are pre-exponential factors endowed with log-normal probability distributions as specified in [11], [12]. These pre-exponential factors correspond to elementary reactions involved in hydrogen oxidation. The resulting marginal probability distribution of a particular fast importance index is shown in Figure 3.

Deterministic CSP simplification employs a user-defined threshold on normalized importance indices to decide which reactions influence the evolution of a particular species, and which reactions may be neglected. The retained reactions enlarge the set of retained species, and the algorithm iterates to convergence. Here, we replace the hard threshold on importance index with a probabilistic criterion. In particular, we consider the conditional value-at-risk (CVaR) [6]. As depicted in Figure 3, the CVaR is the expected value of the importance index given that the importance index exceeds a particular level:

$$\mathrm{CVaR} \equiv \mathbb{E}\left[I_{km} | I_{km} > \mu\right] \tag{3}$$

where μ is determined by $\mathbb{P}\{I_{km} > \mu\} = \alpha$. In this definition, $1-\alpha$ is known as the confidence level while μ is the value-at-risk (VaR). Compared to VaR, numerical evaluation of CVaR is less sensitive to sampling error. Moreover, using CVaR avoids the diffulties of arbitrary threshold selection, and naturally captures the notion that exceeding a threshold by a large amount is undesirable.

Beginning with the importance index distributions calculated above, we use a CVaR threshold in the simplification algorithm to construct simplified mechanisms of various sizes. We then use Monte Carlo simulation with each simplified mechanism to evaluate a probability

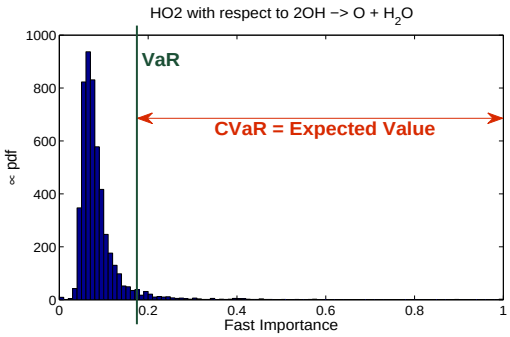

Fig. 3. Probability density function of the fast importance index of $\{2\mathrm{OH} \rightarrow \mathrm{O} + \mathrm{H_2O}\}$ to $\mathrm{HO_2}$. Also shown are the VaR and CVaR at a 95% confidence level.

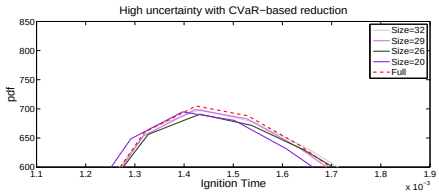

Fig. 4. Probability densities of ignition delay as predicted by simplified mechanisms constructed with risk-based thresholding. Detailed mechanism predictions are depicted with the dashed red line.

density for the predicted ignition delay. Figure 4 shows probability densities resulting from several simplified models (colored solid line), compared to the probability density produced by the full detailed model (dashed line). Agreement is quite good, even for a 20-species mechanism. Figure 5, on the other hand, shows ignition delay probability densities calculated with a different set of simplified mechanisms—mechanisms constructed via the usual deterministic CSP simplification algorithm. In this case, the simplified mechanisms do a much poorer job replicating the output probability density of the detailed model.

ACKNOWLEDGMENT

Y.M. and T.C. acknowledge support from the US Department of Energy (DOE), Office of Advanced Scientific Computing Research (ASCR), Applied Mathematics Program under contract number DE-SC0003564. H.N. also

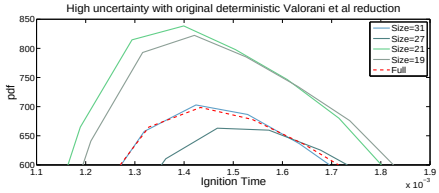

Fig. 5. Probability densities of ignition delay as predicted by simplified mechanisms constructed with the deterministic CSP simplification algorithm. Detailed mechanism predictions are shown with the dashed red line.

acknowledges support from Sandia National Laboratories and from the DOE Office of Basic Energy Sciences (BES) Division of Chemical Sciences, Geosciences, and Biosciences. All three authors gratefully acknowledge support from the 2009 American Recovery and Reinvestment Act. Sandia National Laboratories is a multi-program laboratory operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the US Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

REFERENCES

[1] Lam, S.H., and Goussis, D.A., *Proc. Comb. Inst.*, 22:931–941 (1988).
[2] Maas, U., and Pope, S., *Comb. Flame*, 88:239–264 (1992).
[3] Zagaris, A., Kaper, H.G., and Kaper, T.J., *J. Nonlin. Sci.*, 14:59–91 (2004).
[4] Lam, S.H., *Comb. Sci. Tech.*, 89:375–404 (1993).
[5] Valorani, M., Creta, F., Goussis, D., Lee, J., and Najm, H., *Comb. Flame*, 146:29–51 (2006).
[6] Rockafellar, R. T., and Uryasev, S., *J. Bank. Financ.*, 26:1443–1471 (2002).
[7] Smith, G.P., Golden, D.M., Frenklach, M., Moriarty, N.W., Eiteneer, B., Goldenberg, M., Bowman, C.T., Hanson, R.K., Song, S., Jr., W.C. Gardiner, Lissianski, V.V., and Zhiwei, Q., *www.me.berkeley.edu/gri_mech/.*
[8] Gibbs, A., and Su, F., *Int. Stat. Rev.*, 70:419–435 (2002).
[9] Valorani, M., Creta, F., Donato, F., Najm, H.N., and Goussis, D.A., *Proc. Comb. Inst.*, 31:483–490 (2007).
[10] Najm, H., Debusschere, B., Marzouk, Y., Widmer, S., and LeMaître, O., *Int. J. Numer. Methods Eng.*, 80:789–814 (2009).
[11] Phenix, B.D., Dinaro, J.L., Tatang, M.A., Tester, J.W., Howard, J.B., and McRae, G.J., *Comb. and Flame*, 112:132–146 (1998).
[12] Reagan, M.T., Najm, H.N., Pébay, P.P., Knio, O.M., and Ghanem, R.G., *Int. J. Chem. Kin.*, 37(6):368–382 (2005).